iRepertoire, Inc.

RepSeq & RepSeq+ Data Management Policies

rev.022020a

Table of Contents

I. Overview

iRepertoire, Inc. (herein "iRepertoire") provides products and services for sequencing of the immune repertoire. In addition to these products and services, iRepertoire provides its Customers (herein "customer" or "user") data output (herein "Output") through its web-based analysis platform, iRweb or the iPair Analyzer (herein "Platforms")

II. Research Use Only Data

Data generated from iRepertoire's products and services are for Research Use Only and not to be used for clinical diagnostics.

III. Proprietary Information and Confidentiality

iRepertoire provides products and services for generating immune repertoire sequencing data. iRepertoire does not claim any reach-through rights to sequencing data produced with our technology. Project-specific intellectual property, as it pertains to sequencing data, may be governed by separate material transfer and/or non-disclosure agreements. In these cases, the terms of those agreements supersede this document. All proprietary information (including, without limitation, technologies, methodologies, primer sequences, and protocols) used to produce immune repertoire data remains the sole property of iRepertoire.

IV. Decision-Making and Assessment of Analyzed Data

Any and all decisions or assessments made in relation to patterns or metrics present on iRweb are the sole responsibility of the user. iRepertoire assumes no legal responsibility for decisions based upon or assessments made with Output from Platforms. iRepertoire may provide insight and advice based on publicly-accessible literature and available experimental data upon customer request. These requests may be subject to the bioinformatic support fees detailed in Section VIII and summarized in Table 1.

V. Basic Complimentary Bioinformatics via iRweb and the iPair Analyzer

iRepertoire provides basic bioinformatic analysis complimentary with all services provided. Bioinformatic analyses for data submitted following the use of any of iRepertoire's products is subject to the fees listed in Table 2. Bioinformatic analyses are output to a web-based software, iRweb. iRweb houses analyzed immune repertoire sequencing data that is secured by individual access credentials. The Output within iRweb is accessible only to authorized users for his or her specific analysis. iRepertoire maintains access to all accounts for troubleshooting purposes and customer support only. iRepertoire supports only those features provided in the iRweb platform (listed below). Additional information can be found in iRepertoire's Data Analysis Guide.

- Amino acid lists for unique CDR3s (uCDR3s)

- Diversity measurements, including Shannon Entropy, Diversity Index, and D50

- Unique CDR3 sharing across samples

Data for single-cell projects (iPair/ iPair+) are provided through email, Dropbox®, or a pre-determined location based upon User preference. This data can be viewed using a local application provided by iRepertoire called the iPair Analyzer™. Specific details regarding data output are discussed in the iPair Analyzer User Guide.

VI. Access to Analyzed Data

1. Data Supplied Externally

    1.1. iRepertoire waives liability and responsibility for the long-term or indefinite storage of data produced outside of its facilities with our primer systems. It is the responsibility of the customer to maintain storage, appropriate backup, and access to all data generated in laboratories and on sequencers outside of iRepertoire's facilities. This data may be shared with iRepertoire for the purpose of analyzing data, but iRepertoire has no obligation to store raw sequencing data files for the purpose of a secondary data backup site.

    1.2. Intermediate processed data files such as demultiplexed files, stitched reads, and various stages of the analytic pipeline, are routinely deleted from the server; however, data stored on iRweb remains intact indefinitely. Intermediate files can be shared with customers free-of-charge in the 3 month period following notification of the release of analyzed data. Access to this data must be requested in writing through email or on the Data or Service Submission Form(s). After the initial 3 month period, a fee of $ 25 per gigabyte of data retrieved from long term storage and a flat $250 re-processing fee per flow cell or lane of data will be levied in order access intermediate data after the 3 month's notification time frame.

2. Data Generated Through Service Projects

    2.1. In the event that User libraries are placed in communal pools (e.g., in sequencing runs with other customer libraries or with internal iRepertoire R&D libraries), only demultiplexed data can be provided to customers. This data is sample-specific and extracted via either our proprietary molecular barcoding system or the Illumina® dual-index system, depending upon products utilized in service. This system provides confidentiality and maintains access to data only by authorized parties.

    2.2. In the event that a customer purchases all sequencing space within either a flow cell or lane (as per Illumina® product definitions), raw FASTQ files output by the sequencer can be shared directly with the customer. Sharing can be performed through multiple channels (e.g., Basespace®, FTP/SFTP, Dropbox®, etc.), but methods must be discussed prior to analysis.

    2.3. All MiSeq® runs produced as a part of service are automatically stored within iRepertoire's own Basespace accounts and are therefore subject to all terms and conditions that apply. If the terms and conditions provided for Illumina's Basespace accounts are not agreeable, iRepertoire should be notified prior to sequencing so that the MiSeq instruments can be disconnected from Basespace and stored locally. The latter operation can only be performed when sequencing flow cells are purchased outright on the MiSeq.

    2.4. All HiSeq® run data are shared through either iRepertoire's sftp server or hard drive shipment (not Basespace). Hard drive shipments will incur fees including purchase of the hardware and shipping and handling fees, which are both dependent upon project size. These fees will be established through official quote and PO prior to shipment.

    2.5. Raw (FASTQ files directly from sequencer) or demultiplexed sequencing data is held in temporary local storage for three months. Raw data files are placed in long-term storage both locally and through AWS® cloud storage services.

    2.6. Intermediate processed data files such as demultiplexed files, stitched reads, and various stages of the analytic pipeline, are routinely deleted from the server; however, data stored on iRweb remains intact indefinitely. Intermediate files can be shared with customers free-of-charge in the 3 month period following notification of the release of analyzed data. Access to this data must be requested in writing through email or on the Data or Service Submission Form(s). After the initial 3 month period, a fee of $ 25 per gigabyte of data retrieved from long term storage and a flat $250 re-processing fee per flow cell or lane of data will be levied in order access intermediate data after the 3 month's notification timeframe.

3.

2.7 iRepertoire will store copies of data produced from service projects indefinitely unless removal of such data is specifically requested. Requests to remove raw data from backups will be assessed on a case-by-case basis and will be subject to data handling fees at $150 per hour. For libraries pooled with other projects (without the use of Illumina® dual indices), removal of raw data files is not permissible.

VII. Bioinformatic Support

1. Initial teleconference-based consultations regarding study design (as it relates to read depth and sample coverage), sample requirements, and the initial release of analyzed data are offered as part of our standard service. Bioinformatic consultations are available via teleconference if requests for analyses exceed what is currently offered through the basic data analysis package. Bioinformatic fees are $150 per hour.

2. Included with product and service projects is an optional 1 hour walk-through of the features of iRweb, via teleconference, which is offered in addition to the Data Analysis Guide. One session is included per Principal Investigator's laboratory per year. Additional training is subject to hourly bioinformatic fees.

3. In addition to all of the tools described in section V, iRepertoire provides a 'Raw Data' download in iRweb in which the majority of the underlying data used to generate the graphics in iRweb are available in .csv formats.

4. All graphical portions of the iRweb analyses are accessible to the customer and can be downloaded individually from sample to sample. iRepertoire does not provide mass download of graphics for projects in excess of 10 libraries.

5. High-resolution versions of Tree Maps, graphical representations of CDR3 diversity, can be provided at no cost within 6 months of the notification for release of data. After this 6 month period, Tree Maps can be provided at a cost of $10 per high resolution Tree Map provided.

Table 1. Bioinformatics and Data Management Fee Table	
Raw sequencing data retrieved from long-term (6+ month) storage	$25 per GB
Re-processing fee	$250 per flow cell
Bioinformatic fees	$150 per hour
HD Tree Maps	$10 per image provided

iRepertoire reserves the right to adjust fees at any time for any reason after evaluation of request. These adjustments will be first discussed with User.

VIII. Bioinformatic Analysis of Submitted Data
   1. Our bioinformatic pipelines for RepSeq and RepSeq+ are built around the industry standard FASTQ file format. As our pipeline is built around this file format, the **only two file formats we can accept are FASTQ and BCL*.**

   1. The submission of BCL files requires that the entire Run Folder* is shared with us for processing and the conversion of these files to a FASTQ format **will** incur data management fees. *We do not accept BAM, SAM, FASTA, BED, VCF, or other file format.*

Table 2. Bioinformatic Analysis Fees	Price	Notes
File Manipulations*	$75.00	Includes but not limited to: bcl2fastq and bcl2fastq2 conversion per flow cell or lane; re-compression or concatenation of files; the processing of ≥ 10 demultiplexed libraries.
iSeq 100, MiniSeq, MiSeq (v2 500-cycle), MiSeq (v3 600-cycle)	Complimentary	
HiSeq	$125.00	With the submission of unprocessed FASTQ files
NextSeq Mid-Throughput	$150.00	
NextSeq High-Throughput	$175.00	
NovaSeq	$500.00	

*Please note that iRepertoire reserves the right to define "File Manipulations" as any task beyond the analysis of standard FASTQ-formatted file submissions.

IX. Standard Turnaround Times For Analyses And Priority Analysis
   1. We allow for up to three weeks for the analysis and release of all data analyzed by our pipeline. This is to account for the traffic of analyzing smaller iSeq 100 data sets (≦1 GB in size) to much larger data sets like those produced by the NovaSeq and NextSeq systems (which can exceed 200 GB in total size). Data analysis is currently handled in a "first come, first serve" basis, with the exception of service projects that have paid a Rush fee. The Rush fee for service projects places the customer's data analysis at the head of the current queue, but does <u>not</u> guarantee a given turnaround time.
   2. Priority Analysis fees (see Table 3) can be paid by our customers that have performed sequencing in their own labs, and it is the equivalent of the Rush fee applied to service projects. The analysis of submitted data is placed in the front of the current analysis queue.
   3. In order to be able to provide Priority Analysis, there are very specific rules and procedures that <u>must</u> be adhered to. *<u>Any requirements that are not followed invalidate the Priority Analysis applied to data submissions, regardless of payment.</u>*

(continued on next page)

4. Customers **_must_** send a notification email to DataSupport@iRepertoire.com upon the completion of the upload of their data to our SFTP server or upon the sharing of a link to download the files. Unless a notification email is sent, the Priority Analysis service **_will not_** be applied to the submission.

5. If demultiplexed or stitched files are requested for the Priority submission(s), the notification email sent to DataSupport@iRepertoire.com **_must_** include this request in order to be honored.

6. The Priority Analysis fee **_does not_** guarantee a given turnaround time.

7. Notification of the Priority Analysis **_must occur via email notification_** and **_cannot_** be included in the comments section of Data Submission Forms.

8. Once paid, the Priority Analysis fee is applicable only for 1 calendar year or 365 days from receipt of payment. If the limits of a given tier (next page) are not reached, the Priority Analysis application **_does not_** roll over into the next calendar year.

9. Tiers cannot be extended. Once a limit of a tier is reached, either via submissions or file sizes, a new quote and new tier must be purchased. Customers *may not* pay the difference in order to extend the limits of their Priority Analysis.

10. Unless data is of the RepSeq+ (dam-PCR) format, files uploaded **cannot** be cleaned, processed, or demultiplexed in any way. Any file manipulations by the customer that require extra attention or work for iRepertoire's Data Support team will render the Priority Analysis void for that submission.

11. Priority Analysis will only apply to uploaded or shared FASTQ formatted sequencing files. Priority Analysis **_will not be applicable nor provided for any files other than FASTQ_** (BCL, BAM, FASTA, etc.).

12. Priority Analysis applies to any and all flow cells or lanes that may be included in submissions. If a multi-lane or multi-flow cell platform (HiSeq, NextSeq, NovaSeq) is used and multiple lanes or flow cells are submitted for analysis, each of these flow cells and lanes counts towards the total covered under the Priority Analysis tier.

Table 3.	Priority Analysis Fee	Eample Platforms
2 GB* or 3 flow cells/lanes**	$75.00	iSeq 100
5 GB* or 5 flow cells/lanes**	$75.00	iSeq 100, MiniSeq
10 GB* or 5 flow cells/lanes**	$75.00	iSeq 100, MiniSeq, MiSeq
15 GB* or 3 flow cells/lanes**	$75.00	MiniSeq, MiSeq
20 GB* or 3 flow cells/lanes**	$75.00	MiSeq
25 GB* or 3 flow cells/lanes**	$95.00	MiSeq, HiSeq
50 GB* or 2 flow cells/lanes**	$95.00	MiSeq, HiSeq
60 GB* or 2 flow cells/lanes**	$95.00	HiSeq
75 GB* or 2 flow cells/lanes**	$105.00	HiSeq, NextSeq Mid
100 GB* or 2 flow cells/lanes**	$125.00	NextSeq Mid, NextSeq High
150 GB* or 2 flow cells/lanes**	$150.00	NextSeq Mid, NextSeq High
200 GB* or 2 flow cells/lanes**	$200.00	NextSeq High, NovaSeq

* File size limit is in Gigabytes, not Gigabases.

**Please note that some sequencing platforms contain multiple lanes or flow cells. These may include the the HiSeq, NextSeq, and NovaSeq systems. Each of these flow cells or lanes _does_ count against the total for Priority Analysis.